Technical Report 1114

Enhancing the Efficiency of Tank Gunnery Evaluation: A Strategy Revisited

Joseph D. Hagman
U.S. Army Research Institute

May 2001

United States Army Research Institute for the Behavioral and Social Sciences

U.S. Army Research Institute for the Behavioral and Social Sciences

A Directorate of the U.S. Total Army Personnel Command

EDGAR M. JOHNSON
Director

Technical Review by

Monte D. Smith, L-3 Communications
Bruce Sterling, US Army Research Institute

NOTICES

DISTRIBUTION: Primary distribution of this Technical Report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: TAPC-ARI-PO, 5001 Eisenhower Ave., Alexandria, VA 22333-5600.

FINAL DISPOSITION: This Technical Report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Technical Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) May 2001	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) October 2000 - April 2001

4. TITLE AND SUBTITLE Enhancing the Efficiency of Tank Gunnery Evaluation: A Strategy Revisited	5a. CONTRACT OR GRANT NUMBER
	5b. PROGRAM ELEMENT NUMBER 633007
6. AUTHOR(S) Joseph D. Hagman (U.S. Army Research Institute)	5c. PROJECT NUMBER A792
	5d. TASK NUMBER 207
	5e. WORK UNIT NUMBER H01
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue Alexandria, VA 22333-5600	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. MONITOR ACRONYM ARI
	11. MONITOR REPORT NUMBER Technical Report 1114

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*: Based on the analysis of gunnery scores fired by 171 M1A2 tank crews, an easy-to-use strategy was developed for predicting which crews will, and will not, first-run qualify on Tank Table VIII before all of the typically required ten engagements have been fired. Scores are added as each engagement is fired and the resulting sum is compared to tabular formatted cutoff scores established to support accurate qualification predictions. Adherence to this strategy will help Active Army armor unit commanders to maximize the efficiency of tank gunnery evaluation by reducing the number of first-run engagements fired, as well as the range time and operational tempo (OPTEMPO) resources spent in doing so, by roughly 20% without sacrificing the purpose and intent of the crew-level gunnery certification process.

15. SUBJECT TERMS

Active Army   Armor   Tank Gunnery   Tank Table VIII   Efficiency

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	29	Joseph D. Hagman 208-334-9390

Technical Report 1114

Enhancing the Efficiency of Tank Gunnery Evaluation: A Strategy Revisited

Joseph D. Hagman
U.S. Army Research Institute

Reserve Component Training Research Unit
Ruth H. Phelps, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

May 2001

Army Project Number
20363007A792

Personnel Performance and Training

Approved for public release; distribution unlimited.

FOREWORD

For the Army's armor units to attain and maintain readiness levels in the face of shrinking resources, more efficient ways to conduct tank gunnery training and evaluation must be identified. Although increased reliance on the use of training aids, devices, simulators, and simulations (TADSS) for tank gunnery appears to be paying off, additional efficiencies are also likely to result from efforts to streamline the structure and content of live-fire-based gunnery evaluation exercises, or tables. The rising cost of main gun ammunition, growing restrictions on access to live-fire range/maneuver areas, and the difficulty/cost associated with transporting soldiers/crews to and from these areas suggest that the benefits of more efficient live-fire tank gunnery evaluation could be substantial.

This report describes the results of research showing that the efficiency of live-fire tank gunnery evaluation on Tank Table VIII (the crew-level certification exercise) can be enhanced by changing its content, to include fewer engagements, and its structure, to include performance "gates" to support early qualification and remediation decisions. By making these changes, the Army can save roughly 20% of the ammunition, operational tempo (OPTEMPO), and range time resources normally spent on Tank Table VIII evaluation without jeopardizing its purpose or intent.

This research was conducted by the U.S. Army Research Institute for the Behavioral and Social Sciences Reserve Component Training Research Unit (ARI-RCTRU), whose mission is to improve the effectiveness and efficiency of RC training through use of the latest in training and evaluation technology. This research is supported under Work Package "RC MAX: Maximizing the Resource Efficiency of RC Weapons Qualification" of ARI's Science and Technology Program for Fiscal Year 2000.

This research was conducted to update and extend that initially sponsored by the National Guard Bureau (NGB), under Project SIMITAR (Simulation in Training for Advance Readiness) under a continuing Memorandum of Understanding initially signed 12 June 1985. Findings have been presented to Chief, Combined Arms Training Strategy, U.S. Army Armor School.

Zita M. Simutis
ZITA M. SIMUTIS
Technical Director

Preceding Pages Blank

ENHANCING THE EFFICIENCY OF TANK GUNNERY EVALUATION: A STRATEGY REVISITED

EXECUTIVE SUMMARY

Research Requirement:

Develop a target engagement reduction strategy for enhancing the efficiency of live-fire gunnery evaluation on Tank Table VIII (TTVIII).

Procedure:

The first-run individual engagement and total gunnery scores of 171 M1A2 tank crews undergoing TTVIII qualification firing at Fort Hood, TX, were analyzed via linear regression routines to determine if fewer than the typically required ten engagements can be used to predict successful qualification.

Findings:

From these analyses, an easy-to-use strategy was developed for predicting which armor crews will, and will not, first-run qualify on TTVIII before all ten engagements have been fired. Scores are added as each engagement is fired and the resulting sum is compared to tabular formatted cutoff scores established to support accurate qualification predictions.

Use of Findings:

Adherence to this strategy will help Active Army armor unit commanders to maximize the efficiency of tank gunnery evaluation by reducing the number of first-run engagements fired, as well as the range time and operational tempo (OPTEMPO) resources spent in doing so, by roughly 20% without sacrificing the purpose and intent of the crew gunnery certification process.

ENHANCING THE EFFICIENCY OF TANK GUNNERY EVALUATION: A STRATEGY REVISITED

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Enhancing the Efficiency of Tank Gunnery Evaluation: A Strategy Revisited

Introduction

For the Army's combat units to attain and maintain required readiness levels in the face of shrinking resources, more efficient ways to conduct crew-served weapons training and evaluation must be identified. Although it is not always clear how to train and evaluate more efficiently, various approaches based on the use of training aids, devices, simulators, and simulations (TADSS) appear to be paying off. The Army National Guard's (ARNG's) decision to increase TADSS usage in an effort to improve the efficiency of tank gunnery training and evaluation (e.g., Krug & Pickell, 1996, February), for example, has resulted in the development of a number of useful TADSS-based products. These include a tool for predicting crew-level, live-fire tank gunnery performance from that fired on the Conduct-of-Fire Trainer (COFT) (Hagman & Smith, 1996), a strategy for using this tool in support of home station, TADSS-based gunnery training and evaluation (Hagman & Morrison, 1996), and a host of other TADSS-based strategies designed to maximize the payoff from resources spent (e.g., Shaler, 1994; U.S. Army Armor School, 1995).

Although increased reliance on the use of TADSS is likely to produce continued improvements in the efficiency of tank gunnery training and evaluation in the Active, as well as the Reserve, Component (AC/RC), additional efficiencies are also likely to result from efforts to streamline the structure and content of live-fire-based gunnery evaluation exercises (i.e., tables). The rising cost of main gun ammunition, growing restrictions on access to live-fire range/maneuver areas, and the difficulty/cost associated with transporting soldiers/crews to and from these areas suggest that the benefits of more efficient live-fire tank gunnery evaluation could be substantial.

To this end, the U.S. Army Research Institute (ARI) (i.e., Smith & Hagman, 1998; Hagman & Smith, 1999, March-April) has developed a strategy for enhancing efficiency by reducing the number of target engagements fired on Tank Table VIII (TTVIII) (i.e., the crew-level proficiency certification exercise) without compromising evaluative validity. TTVIII consists of ten engagements, selected from a set of twelve, that encompass a variety of offensive and defensive combat scenarios involving both single and multiple stationary and moving targets (Department of the Army, 1993). Although performance demonstrated on a ten-engagement scenario has historically been used to certify crew-level gunnery proficiency on TTVIII, the tightening of resources over the last decade prompted ARI's development of a strategy that employs fewer engagements to do the same job.

In general, this target engagement reduction strategy uses the crew performance scores obtained as each engagement is fired to predict what the final ten-engagement score would be. These predictions are then used to either qualify a crew or send it back for remedial training (e.g., using TADSS) -- two actions that heretofore have had to await the firing of all ten engagements. The earlier in the TTVIII target engagement firing

sequence that these predictions can be made, the greater will be the resource (e.g., range time, ammunition, operational tempo [OPTEMPO]) savings.

Recently, the TTVIII engagements upon which this strategy is predicated have been changed (Department of the Army, 1998). Consequently, the strategy needs to be reexamined to determine its applicability to this new set of engagements. The present research examines this applicability in regard to AC tank crew performance, describes the process followed in doing so, and concludes with an estimate of resource savings to be expected from strategy usage.

Method

Data

The first-run, individual engagement and total gunnery scores of 171 M1A2 tank crews undergoing semiannual TTVIII qualification firing at Fort Hood, TX, were analyzed. Total scores could vary from 0-1000 (i.e., up to 100 points for each of ten individual engagements) with a qualification score of 700 or more needed for crew proficiency certification.

Analyses

Following procedures described by Smith and Hagman (1998), stepwise multiple regression routines (SPSS, 1994) were used to determine if fewer than ten TTVIII engagements (i.e., subsets) could be used to predict total scores (i.e., those based on the firing of all ten engagements). The best subsets of from one to nine engagements were identified and the predictive validity specified for each. Subset identification was based on part-whole Pearson Product-Moment coefficients of correlation (r) between individual engagement and total scores. The best individual predictor (i.e., engagement score) was then used to construct a prediction equation of the form:

$$Y' = B_0 + B_1(X_1) \qquad (1)$$

Where Y' was the predicted total score, B_0 was the intercept/constant (or theoretical total score when the predictor variable equals zero), B_1 was the empirically derived regression coefficient linking changes in the criterion variable (i.e., total score) with changes in the predictor variable (i.e., engagement score), and X_1 was the engagement score most highly correlated with the criterion variable.

The remaining nine engagements were then examined to identify the one that best enhanced the predictive power of the first engagement. After all pairwise combinations of the original predictor with each of the remaining potential predictors were tested and the best second predictor identified, a new multiple regression prediction equation was developed using the combined predictive power of the two best predictors. The new prediction equation took the following form:

$$Y' = B_0 + B_1(X_1) + B_2(X_2) \quad (2)$$

Where Y', B_0, B_1, and X_1, were as defined in Equation 1, B_2 was the empirically derived regression coefficient linking changes in the criterion variable with changes in the second predictor variable, and X_2 was the second predictor variable--the one that best augmented the predictive power of the original predictor.

The two-predictor multiple regression prediction equation was then fitted to the data, yielding a new set of criterion residual scores which represented the criterion scores after the linear effect of the first two predictors was removed. The remaining engagements were then examined to identify the one that best enhanced predictive power when it was added to the two-predictor model to form a three-predictor model. This step produced a new prediction equation structurally similar to Equation 2 except that it contained the term, $B_3(X_3)$, which represented the third predictor and its empirically determined regression coefficient:

$$Y' = B_0 + B_1(X_1) + B_2(X_2) + B_3(X_3) \quad (3)$$

This procedure was repeated as long as additional predictor variables significantly enhanced the predictive power of the resulting equation, which took the general form:

$$Y' = B_0 + B_1(X_1) + B_2(X_2) + B_3(X_3) + B_n(X_n) \quad (4)$$

Where Y', B_0, B_1, X_1, $B2$, X_2, B_3, and X_3 were as defined in Equations 1-3 and the term $B_n(X_n)$ represented the nth predictor variable (X_n) and its empirically determined regression coefficient, B_n.

Results

Descriptive Data

Table 1 shows the means and standard deviations (*SDs*) for total and individual engagement scores while Table 2 shows the part-whole correlations obtained between the total score and each engagement score along with the intercorrelations obtained between each engagement pair. The part-whole correlations ranged from .18 (Engagement B4) to .44 (Engagement B1), with a mean of .32, and were all statistically significant with the rejection region for this and all subsequent analyses set at .05. In contrast, the intercorrelations among individual engagements ranged from -.14 (Engagements A1/A4 and A4/B4) to .14 (Engagements A4/B1), with a mean of .01, and were all nonsignificant. The low intercorrelations among individual engagements revealed that performance on one engagement cannot be predicted on the basis of performance on another, whereas the higher part-whole correlations revealed that every engagement had the potential of making a contribution to total score predictions.

Table 1
TTVIII Descriptive Data

Engagement	Mean	SD
Total	719.84	90.10
A1	73.32	29.95
A2S	58.33	30.14
A3	68.46	29.59
A4	89.85	26.98
A5/A5A	73.73	25.79
B1	64.60	38.54
B2B2A	65.26	27.85
B3S	68.62	21.94
B4	80.50	24.46
B5	80.16	20.38

Table 2
TTVIII Correlation Matrix

	A1	A2S	A3	A4	A5/A5A	B1	B2/B2A	B3S	B4	B5
Total	.20*	.39*	.35*	.28*	.36*	.44*	.39*	.34*	.18*	.28*
A1		-.09	-.04	-.14	.01	-.08	-.03	.03	-.02	.12
A2S			.01	.05	-.03	-.07	.11	.12	.07	.12
A3				.03	.06	-.02	-.03	.04	-.01	.06
A4					-.08	.14	-.06	.09	-.14	.02
A5/A5A						.07	.03	.13	.03	.02
B1							.10	-.01	-.11	.03
B2/B2A								.02	.03	.04
B3S									-.09	-.05
B4										-.05

$^*p \leq .05$

Cross Validation

Before proceeding with additional analyses, a split-half, cross-validation design (Tatsuoka, 1969) was used to examine the potential generalizability of the present performance scores to those of future AC tank crews. Eighty-six of the 171 tank crews in the current sample were assigned at random to the normative group while the remaining eighty-five crews were assigned to the cross-validation group. Stepwise, least-squares, multiple regression routines were then used to identify the best predictive subsets of from 1 through 9 predictor variables for the normative group with a separate equation developed for each subset size. All prediction equations were significant, producing Multiple *R's* ranging from *.50* (based on 1 predictor) to *.98* (based on 9 predictors) and *F-ratios* ranging from *27.95 (1, 84)* to *194.74 (9, 76)*.

The generalizability/validity of the normative group equations was then tested on the cross-validation group and the accuracy of predictions for the two groups compared. The resulting z tests for differences between Multiple R's (Hays, 1963, p. 532) revealed that, regardless of the number of predictors involved, the predictions developed from

normative group data accounted for a comparable amount of total score variance in the cross-validation group. Thus, the predictions were found to be valid and, therefore, likely to maintain similar accuracy levels when used to predict the TTVIII total scores of future AC tank crew samples. Given the similar outcomes of the separate group analyses, along with the desire to obtain the best possible predictions from the largest sample size possible, subsequent analyses were performed on pooled-group data.

Pooled-Group Prediction Equations

Pooled-group prediction equations were developed for the best predictive subsets of 1 through 9 engagements. The order of engagement entry into the equations is shown in the first column of Table 3. The equations themselves are shown in Table 4.

Table 3
Stepwise Multiple Regression Results

Order of Entry	Multiple R	Adjusted R^2	SE	df	F
1. B1	.44	.19	81.10	1, 169	40.85*
2. A2S	.61	.37	71.54	2, 168	50.84*
3. A3	.71	.49	64.26	3, 167	55.73*
4. A5/A5A	.78	.59	57.59	4, 166	65.52*
5. B2/B2A	.83	.68	51.10	5, 165	72.72*
6. A1	.88	.76	44.02	6, 164	91.38*
7. A4	.92	.84	35.85	7, 163	130.12*
8. B4	.95	.90	28.89	8, 162	186.39*
9. B3S	.98	.95	20.60	9, 161	343.47*

* $p \leq .05$

The prediction equation for each subset size was significant, producing Multiple *R's* ranging from *.44* (based on one predictor) to *.98* (based on nine predictors) and *F- ratios* ranging from *40.85 to 343.47.* The first predictor to enter the equation (Engagement B1) had the highest zero-order correlation (.44) with the criterion and accounted for almost one fifth (Adjusted [for shrinkage] R^2 = 19%) of TTVIII total score variation. The addition of the second predictor (Engagement A2S) boosted the proportion of explained variance to 37%, with this proportion steadily increasing with the addition of each subsequent predictor until a 95% total score predictive accuracy was achieved from the firing of the nine-member subset which included all engagements except for B5. Thus, after firing these nine engagements, total score predictions (i.e., the score obtained if all ten engagements were fired) would have an accuracy rate of 95%. After firing the first eight engagements, the predictive accuracy rate would be 90%, and so forth.

Although the firing order of these engagements would be limited only by practical considerations (e.g., firing the day engagements first and then the night engagements, or vice versa), implementation of the above prediction approach would saddle the user with having to complete a cumbersome calculation procedure in order to arrive at the desired TTVIII total score predictions. The commander who wants to trim two engagements from the standard ten, for instance, must have his crews fire eight engagements, score

them, multiply each score by its respective regression coefficient shown in Table 4 (e.g., 1.01 for Engagement A1, 1.18 for Engagement A2S), and then add the prediction equation constant (i.e., 162.74) to arrive at the predicted total score for each crew. This procedure would be time-consuming and subject to error when performed on the range.

Table 4
Prediction Equations for Subset Sizes 1 to 9

Subset Size	Prediction Equation
1	Y' = 653.21 + 1.03 (B1)
2	Y' = 573.75 + 1.11(B1) + 1.28(A2S)
3	Y' = 500.32 + 1.12(B1) + 1.27(A2S) + 1.07(A3)
4	Y' = 423.87 + 1.07(B1) + 1.29(A2S) + 1.01(A3) + 1.11(A5/A5A)
5	Y' = 372.19 + 1.00(B1) + 1.18(A2S) + 1.04(A3) + 1.10(A5/A5A) + .97(B2/B2A)
6	Y' = 298.06 + 1.05(B1) + 1.27(A2S) + 1.08(A3) + 1.07(A5/A5A) + .97(B2/B2A) + .87(A1)
7	Y' = 206.02 + .95(B1) + 1.22(A2S) + 1.05(A3) + 1.16(A5/A5A) + 1.05(B2/B2A) + .98(A1) + .96(A4)
8	Y' = 162.74 + 1.00(B1) + 1.18(A2S) + 1.06(A3) + 1.13(A5/A5A) + 1.03(B2/B2A) + 1.01(A1) + 1.07(A4) + .87(B4)
9	Y' = 79.10 + 1.01(B1) + 1.09(A2S) + 1.04(A3) + 1.02(A5/A5A) + 1.02(B2/B2A) + .97(A1) + 1.0(A4) + .95(B4) + .93(B3S)

Shortcut Approach

Both of these potential negative outcomes raise the question of whether an easier-to-implement (i.e., shortcut) prediction procedure could be used on the range with minimal, if any, sacrifice of predictive accuracy. Visual inspection of the regression coefficients shown in Table 4 revealed that most values were close to 1.00, thereby suggesting the possibility of eliminating them altogether and substituting a procedure that weighs each engagement equally and eliminates the constant. If so, a shortcut prediction procedure could be reduced to three simple steps:

1: Add the engagement scores of the desired engagement subset size.
2: Divide the sum by N_{sub}, the number of engagements in the subset.
3: Multiply the quotient by 10.

Accordingly, each engagement is weighted equally by dividing by the number of engagements in the subset (N_{sub}), and the mean of all engagements in the subset is extrapolated to a ten-engagement total score (by multiplying by 10), thereby lumping the variance from all available engagements into a single predictor.

The efficacy of this procedure was tested by constructing a series of shortcut predictor variables; one for each subset size, based on the best predictive set of engagements identified in the stepwise regression procedures. For the two-engagement subset, for example, the first shortcut predictor variable was calculated by the procedure: [(B1 + A2S)/2] x 10. Thus, if a crew fired a score of 55 on Engagement B1 and a score

of 97 on Engagement A2S, its predicted shortcut total score would be 760. This shortcut score could then used to predict a crew's total scores.

The results of the shortcut test are shown in Table 5. The column under "Regression Models" shows Adjusted R^2 values for each subset size for the best engagement predictors, as determined by stepwise multiple regression procedures. The column under "Shortcut Models" shows Adjusted R^2 values using analogous shortcut regression procedures based on the same best predictors for each subset size. The values in the two columns turned out to be identical and indicate that the shortcut prediction method can indeed be used successfully with reduced subsets of any size.

Table 5

Adjusted R^2 Values for Full Regression Models vs Shortcut Regression Models

Subset Size	Regression Models Adjusted R^2	Shortcut Models Adjusted R^2
2	.37	.37
3	.49	.49
4	.59	.59
5	.68	.68
6	.76	.76
7	.84	.84
8	.90	.90
9	.95	.95

Random Engagement Subsets

Even with the easier calculational procedures afforded by the shortcut method, using the statistically identified best predictors requires that specific engagements be fired for each subset (i.e., B1 and A2S for the two-engagement subset) and introduces the possibility of "training to the test" in order to save time, especially if a commander were to select a TTVIII certification subset with relatively few engagements. To encourage training on the widest variety of engagements possible in preparation for TTVIII firing, engagements to be included in any particular subset could be selected at random. This would necessitate training on all possible engagements because crews would not know beforehand which ones would be included in the subset selected for certification.

To determine if engagements can indeed be selected at random without appreciably degrading predictive accuracy, random subsets of engagements were constituted for subset sizes ranging from two to nine. Randomization was accomplished by labeling ten coins (i.e., one for each engagement). The coins were then placed in a hat and blindly drawn without replacement to create a random subset of engagements of the desired size. Once a subset was created, drawn coins were returned to the hat, all coins were again shaken to redistribute them inside the hat, and the process was repeated until a total of five random subsets were created for each subset size from two to nine. Subsets of size six or greater were created by random exclusion. That is, to create a six-engagement

subset, four engagements were drawn randomly and excluded. The six engagements remaining in the hat became the subset. For subsets of size seven, three engagements were randomly excluded, and so on. This produced five 2-engagement random subsets, five 3-engagement random subsets, and so on, up to and including five 9-engagement random subsets. In all, 40 random subsets were constructed, five for each of the eight possible subset sizes.

Multiple regression procedures based on the shortcut method were used to construct a prediction equation for each of the 40 random subsets created (See Appendix A for the specific engagements drawn for each random subset and associated statistical information.) The predictive power of the random engagement subsets was then compared to that of the best engagement subsets. As expected, the best predictors accounted for a greater proportion of total score variance than that accounted for by the mean random engagement predictors (See Table 6) across the different subset sizes. Although predictive accuracy uniformly favored the former over the latter predictors, the mean difference between the two was only 12% across the eight subset sizes. Arguably, this difference is more than offset by the elimination of the possibility of training to the test when random engagement subsets are used. In addition, use of these subsets affords armor unit commanders maximum flexibility in the selection of engagements that can now be best fitted to existing targeting sequence layouts on the range, thereby promoting the most efficient use of OPTEMPO resources set aside for TTVIII firing. Consequently, the remaining sections of this report assumes that engagements are selected on a random basis.

Table 6.
Adjusted R^2 and SE Values for Random and Best Engagement Subset Sizes 2-9

Subset	Random		Best	
Size	Adjusted R^2	*SE*	Adjusted R^2	*SE*
2	.23	79.01	.37	71.46
3	.30	75.44	.49	64.03
4	.41	69.21	.59	57.45
5	.53	61.47	.68	50.75
6	.60	57.03	.76	44.28
7	.72	47.70	.84	36.06
8	.82	37.25	.90	29.00
9	.89	29.12	.95	20.40

Predicting Qualification

This section extends the analytical procedures outlined thus far to the prediction of crew qualification status. The capability to predict which crews will, and will not, first-run qualify, as each engagement is fired, would allow armor unit commanders to award early qualification status to some crews (i.e., those predicted to fire 700 or above) and to identify others in need of remediation (i.e., those predicted to fire below 700). The specific procedures for deriving such predictions are described below.

Early elimination predictions. For any given random engagement subset size (N_{sub}), the minimum cutoff score (E_{Elim}) necessary to avoid early elimination (i.e., removal from the range for remediation purposes) can be predicted from the general equation:

$$E_{Elim} = ([700 - (1.65 * SE_{ind})] / 10)* N_{sub} \qquad (5)$$

Where E_{Elim} is the minimum subset cutoff score needed to avoid early elimination, 700 is the minimum ten-engagement-based score required for qualification, 1.65 is the normal deviate (in a one-tailed directional test) for a 95% confidence interval (lower bound), SE_{ind} is the standard error of estimate for the minimum cutoff score of 700 (assuming that the actual probability of firing 700 will follow a normal distribution), and N_{sub} is the engagement subset size on which the prediction is based. The SE_{ind} values used in this case to test the prediction for each random subset size represent the mean SE_{ind} obtained from the five randomly formed subsets identified earlier for each subset size (See Appendix A for the engagement composition of each subset.).

Testing early elimination predictions. The early elimination predictive cutoff values, shown in Table 7, were then tested against the gunnery scores fired by the pooled group sample. For each subset size of engagements, tank crews' scores were compared to the cutoff scores in the far right column. Crews with scores that were less than the corresponding tabled cutoff score were predicted to have less than a 5% chance of first–run qualification (i.e., Q1), if they were to proceed with the firing of all ten engagements.

Table 7
Minimum E_{Elim} Values to Avoid Early Elimination

Subset Size	Prediction Equation	Minimum E_{Elim}
2	E_{Elim} = [700 - (1.65 x 79.41)]/10 x 2	114
3	E_{Elim} = [700 - (1.65 x 75.82)]/10 x 3	172
4	E_{Elim} = [700 - (1.65 x 69.56)]/10 x 4	234
5	E_{Elim} = [700 - (1.65 x 61.78)]/10 x 5	299
6	E_{Elim} = [700 - (1.65 x 57.32)]/10 x 6	363
7	E_{Elim} = [700 - (1.65 x 47.94)]/10 x 7	435
8	E_{Elim} = [700 - (1.65 x 37.44)]/10 x 8	511
9	E_{Elim} = [700 - (1.65 x 29.27)]/10 x 9	587[a]

[a] Mathematically eliminated with a score < 600

The early elimination test results are shown in Table 8. Based on the scores obtained for a subset containing two engagements (See first data row.), for example, 25 crews (14.7% of the total sample) were identified for early elimination with 100% accuracy, as shown in the last column. Thus, all crews predicted to be nonqualifiers after firing two engagements indeed did not qualify after firing ten (referred to as a Hit).

Table 8

A Test of Early Elimination Predictions

Crews	Subset Size	Cutoff Score (≤)	Identified Crews: # of Crews	Cum #	Cum %	Hits	Predictive Accuracy %
171	2	114	25	25	14.7	25	100
146	3	172	6	31	18.1	6	100
140	4	234	4	35	20.5	4	100
136	5	299	1	36	21.1	1	100
135	6	363	2	38	22.2	2	100
133	7	435	2	40	23.4	2	100
131	8	511	3	43	25.1	3	100
128	9	587[a]	1	44	25.7	1	100

Note. Cum = Cumulative

[a] Mathematically eliminated with a score < 600.

Once these 25 crews were removed from the analysis, the three-engagement subset size prediction was tested on the remaining 146 crews. Based on this test, six additional crews (yielding a cumulative 18.1% of the total sample) were identified for early elimination. All six of these crews indeed failed to Q1. Finally, by the time subset size reached nine (See bottom row.), 25.7% of the total sample (44 crews) had been correctly predicted as unlikely to Q1.

Early qualification predictions. An adaptation of Equation 5 can be used to identify crews likely to fire Q1:

$$E_{\text{Qual}} = ([700 + (1.65 * SE_{ind})] / 10)* N_{\text{sub}} \qquad (6)$$

Where E_{Qual} is the minimum cutoff score necessary for early qualification, and all other terms are as defined for Equation 5. Crews scoring at or above the specified subset cutoff scores could be awarded early Q1 status with 95% confidence that, had they been allowed to fire all ten engagements, they would have received a score of 700 or more.

Table 9 shows the required E_{qual} cutoff score for each subset size. After completing the number of engagements listed in the Subset Size column, crews achieving a cumulative score equal to or greater than the corresponding cutoff value in the Minimum E_{qual} column would be eligible for early Q1.

Testing early qualification predictions. The same steps used for testing early elimination were also used for testing early qualification. The results of the latter are shown in Table 10. For the two-engagement subset, 18 crews (10.5% of the total sample) were predicted to Q1. This prediction turned out to be correct for all 18 crews, for a hit rate accuracy of 100%. Once these 18 crews were removed from the analysis, the prediction for the three-engagement subset size was tested on the remaining 153 crews. Here, 11 additional crews (yielding a cumulative 17% of the total sample) were identified

for early Q1, again with 100% predictive accuracy. Subsequent rows may be interpreted in a similar manner with each row containing the data for its respective subset size.

Table 9

Minimum E_{qual} Values for Early Qualification

Subset Size	Prediction Equation	Minimum E_{qual}
2	$E_{qual} = [700 + (1.65 * 79.41)] / 10 * 2$	166
3	$E_{qual} = [700 + (1.65 * 75.82)] / 10 * 3$	248
4	$E_{qual} = [700 + (1.65 * 69.56)] / 10 * 4$	326
5	$E_{qual} = [700 + (1.65 * 61.78)] / 10 * 5$	401
6	$E_{qual} = [700 + (1.65 * 57.32)] / 10 * 6$	477
7	$E_{qual} = [700 + (1.65 * 47.94)] / 10 * 7$	545
8	$E_{qual} = [700 + (1.65 * 37.44)] / 10 * 8$	609
9	$E_{qual} = [700 + (1.65 * 29.27)] / 10 * 9$	673

Table 10

A Test of Early Qualification Predictions

			Identified Cases				
Crews	Subset Size	Cutoff Score (≥)	# of Crews	Cum #	Cum %	Hits	Predictive Accuracy %
171	2	166	18	18	10.5	18	100
153	3	248	11	29	17.0	11	100
142	4	326	4	33	19.3	4	100
138	5	401	4	37	21.6	4	100
134	6	477	5	42	24.6	5	100
129	7	545	2	44	25.7	2	100
127	8	609	13	57	33.3	13	100
114	9	673	10	67	39.2	10	100

Note. Cum = Cumulative

Combined effects of early elimination and early qualification predictions. Based on the combined early elimination and qualification predictions shown below in Table 11, about a quarter of the crews (See the last column.) in the present sample could have been removed from the range after firing only two engagements. Over one third could have been removed after firing only three engagements, about one half after seven engagements, and so on. Finally, the bottom row shows that less than 35% of the crews would have needed to fire all ten engagements to determine their Q1 status, because their scores continued to fall between the cutoff values established for early elimination and qualification.

Table 11
Combined Early Elimination and Early Qualification Predictions

Subset Size	Early Elimination Cutoff (<)	Early Qualification Cutoff (≥)	Elimination %	Qualification %	Cumulative %
2	114	166	14.7	10.5	25.2
3	172	248	3.5	6.4	35.0
4	234	326	2.3	2.3	39.6
5	299	401	.6	2.3	42.5
6	363	477	1.2	2.9	46.6
7	435	545	1.2	1.2	49.9
8	511	609	1.8	7.6	58.4
9	587[a]	673	1.2	5.8	65.4

[a] Mathematically eliminated with a score < 600.

Implementing the Strategy

Figure 1 shows how the proposed target engagement reduction strategy would be implemented in an armor battalion using the cutoff scores shown in the second and third columns of Table 11. In general, once the unit commander has randomly selected the ten TTVIII engagements (from the set of 12 possible) to be fired, crew gunnery proficiency would be evaluated after the firing of each engagement, irrespective of firing order.

All crews would begin firing with the first two engagements. Those not scoring at least 114 would be pulled from the range and given TADSS-based remedial training, perhaps on the COFT or Abrams Full-Crew Interactive Simulation Trainer (AFIST). Once their TADSS-based performance suggests a reasonable probability of successful live-fire TTVIII qualification (See Hagman & Smith, 1996, for the description of a COFT-based tool for predicting such), they would be given one rerun attempt, starting from the beginning with the first two engagements.

First-run crews scoring 166 or higher after firing the first two engagements would be awarded early qualification (i.e., Q1e); those scoring from 114 to 165 would go on to the third engagement. Crews not scoring at least 172 after three engagements would undergo remediation before beginning their rerun from the beginning. Rerun crews would be evaluated as if they were firing their first run, except that predictions would now apply to Q2 rather than Q1. Those predicted to need remediation as a result of low scores on their rerun would receive an unqualified rating. First-run crews scoring 248 or higher after three engagements would be awarded early qualification; those scoring from 172 to 247 would go on to the fourth engagement, and so on.

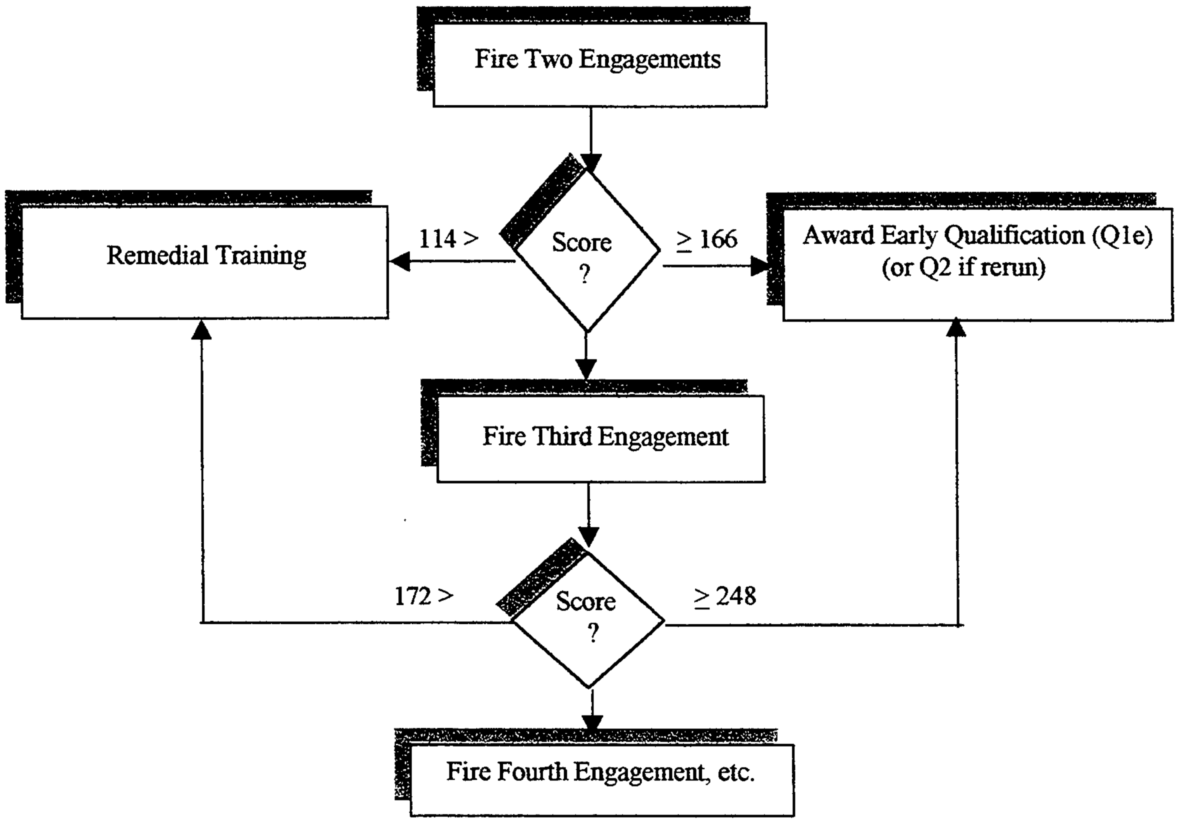

Figure 1. Flowchart of TTVIII engagement sequence.

Resource Savings

Generally speaking, the earlier in the target engagement firing sequence that early elimination or qualification predictions can be made, the greater the resource savings will be. Assuming that each engagement accounts for roughly 10% of the total resources spent on TTVIII, crews predicted to Q1 after only two engagements would save about 80% of the resources needed to fire all ten. Those predicted to Q1 after three engagements would save 70%, and so on.

Resources are likely to be saved by crews predicted to early Q1 as well as by those predicted to need remediation. Using the results obtained with pooled group data, the number of crews in a 44-crew armor battalion that would be predicted to Q1 after each engagement, as well as the predicted number of engagements they would save, was calculated. As shown in Table 12, five crews would be predicted to Q1 after two engagements (i.e., 10.5% [Table 5] x 44 crews in the battalion) and save a total of 40 engagements (i.e., 5 crews x 8 engagements), three crews would be predicted to Q1 after three engagements and save 21 engagements, and so on, with 88 engagements saved in all by the entire battalion. Thus, on crews predicted to early Q1 alone, 20 % (88/440) of an armor battalion's first-run engagements could be saved merely by applying the above strategy.

Table 12

Predicted # of Engagements Saved by an Armor Battalion on the First Run of TTVIII

# of Engagements Fired	Predicted # of Early Q1 Crews	Predicted # of Engagements Saved
2	5	40
3	3	21
4	1	6
5	1	5
6	1	4
7	1	3
8	3	6
9	3	3
	Total: 18	Total: 88

Battalion resources should also be saved on crews predicted to need remedial training simply because they can be identified before they have fired all ten TTVIII engagements. How much savings, however, would depend on how many rerun engagements are fired once these crews are returned to the range. Having these crews start their reruns from the beginning, and then reapplying the engagement subset cutoff scores, should help to maximize the savings on their rerun attempt. Thus, in general, reducing the number of engagements fired through early prediction of which crews will, and which will not, first-run qualify should translate into fewer main gun rounds, less range time, and reduced OPTEMPO costs each year on TTVIII.

Summary and Conclusions

The findings of the present research reinforce and extend the findings of that performed previously (Smith & Hagman, 1998; Hagman & Smith, 1999, March-April) suggesting that more efficient live-fire gunnery evaluation is possible in AC armor units simply by changing current TTVIII content, to include fewer engagements, as well as structure, to include performance cutoff scores or "gates" to support early qualification and remediation decisions. To this end, the target engagement reduction strategy provided herein will promote TTVIII efficiency by reducing roughly 20% of the number of engagements fired, as well as the OPTEMPO resources and range time spent in doing so, all without sacrificing the purpose and intent of the evaluation process. These savings can then be used to offset future resource cuts, pocketed, or used for other purposes such as platoon-level gunnery (e.g., TTXII).

This strategy, provided in flowchart format, can be easily implemented at the company or battalion level by first adding the scores of engagements as they are fired and then determining the resulting summed score status in relation to early elimination and qualification cutoff scores provided in accompanying tabular format. To the extent that future tank crew performance is similar to that found in the present research, the resulting qualification vs nonqualification predictions will, on the average, be accurate 95% of the time. Although the validity of these predictions was tested herein using split-half cross-

validation procedures, those interested in performing further tests of predictive validity should consider doing so with TTVIII data gathered from several different tank crew samples, just to be on the safe side.

Although the strategy presented here is specific to the AC, it should also apply to the RC with the exception of the tabled early elimination and qualification cutoff score values. Additional data still need to be collected from RC tank crews in order to determine exactly what these values would be.

References

Department of the Army (1993). *Tank gunnery training (Abrams) (FM 17-12-1-2).* Washington, DC: Author.

Department of the Army (1998). *Tank gunnery training (Abrams) (FM 17-12-1-2).* Washington, DC: Author.

Hagman, J. D., & Morrison, J. E. (1996, November-December). A strategy for efficient device-based tank gunnery training in the Army National Guard. *Armor*, 48-50

Hagman, J. D., & Smith, M. D. (1996). Device-Based prediction of tank gunnery performance. *Military Psychology, 8,* 59-68.

Hagman, J. D., & Smith, M. D. (1999, March-April). How the guard could cut costs on Table VIII without really trying. *Armor, CVIII(2),* 47-49.

Hays, W. L. (1963). *Statistics.* New York: Holt, Rinehart, and Winston

Krug, R. E., & Pickell, G. A. (1996, February). SIMITAR sharpens the guard. *Army*, 57-59.

Shaler, M. D. (1995). *Compressed gunnery program for Abrams Tank battalions: Trainer's handbook.* Arlington, VA: Houston Associates.

Smith, M. D., & Hagman, J. D. (1998). *Enhancing the resource efficiency of live-fire tank gunnery evaluation* (Tech. Rep. 1088). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A368 641)

SPSS, Inc. (1994). *SPSS 6.1 for Windows update.* Chicago, IL: SPSS, Inc.

Tatsuoka, M.M. (1969). *Selected topics in advanced statistics: Validation studies.* Champaign, IL: The Institute for Personality and Ability Testing.

U.S. Army Armor School (1995). *Enhanced mounted brigade training strategy.* Fort Knox, KY: Author.

Appendix A
Random Engagement Subsets

Two-Engagement random subset. The first five rows of Table A-1 present the shortcut-method-based multiple regression results for the five randomly drawn subsets containing two engagements. Means in the sixth line of the table are based on these subsets. The last line in the table provides analogous multiple regression results based on the two best predictors (Engagements B1 and A2S). Adjusted R^2s for the random subsets ranged from .16 to .37 and accounted for an average of 23% of criterion (total TTVIII score) variance and produced an average *SE* of 79.01 along with significant *F*-values in excess of 33. By comparison, the two best predictors accounted for 37% of adjusted criterion variance and had an *F*-value of 101.30.

Table A-1
Two-Engagement Random Subsets vs The Two Best Predictors

Engagements	Multiple R	Adjusted R^2	$F(1,169)$	SE
A1, B1	.49	.23	52.58*	78.92
A3, B3S	.47	.22	48.13*	79.73
B2/B2A, B4	.40	.16	33.11*	82.64
B1, A2S	.61	.37	101.30*	71.46
A4, B3S	.41	.17	34.82*	82.29
Mean Random	.48	.23	53.99	79.01
Best Two	.61	.37	101.30*	71.46

* $p \leq .05$ for this and all following tables

Three-Engagement random subsets. The first five rows of Table A-2 present the shortcut-method-based multiple regression results for the five randomly drawn subsets containing three engagements. Means in the sixth line of the table are based on these subsets. The last line in the table provides analogous multiple regression results based on the three best predictors (Engagements B1, A2S, and A3). Adjusted R^2s for the random subsets ranged from .22 to .40 and accounted for an average of 30% of criterion variance and produced an average *SE* of 75.44 along with significant *F*-values in excess of 48. By comparison, the two best predictors accounted for 49% of adjusted criterion variance and had an *F*-value of 167.57.

Four-Engagement random subsets. The first five rows of Table A-3 present the shortcut-method-based multiple regression results for the five randomly drawn subsets containing four engagements. Means in the sixth line of the table are based on these subsets. The last line in the table provides analogous multiple regression results based on the four best predictors (Engagements B1, A2S, A3, and A5/A5A). Adjusted R^2s for the random subsets ranged from .36 to .46 and accounted for an average of 41% of criterion variance and produced an average *SE* of 69.21 along with significant *F*-values in excess of 94. By comparison, the two best predictors accounted for 59% of adjusted criterion variance and had an F-value of 249.22.

Table A-2
Three-Engagement Random Subsets vs The Three Best Predictors

Engagements	Multiple R	Adjusted R^2	$F(1, 169)$	SE
B4, A5/A5A, B5	.47	.22	48.30*	79.69
B4, A4, B3S	.48	.23	51.42*	79.13
A3, B5, A2S	.56	.31	78.17*	74.73
A2S, A5/A5A, B2/B2A	.63	.40	112.19*	70.06
A1, B1, B3S	.58	.33	85.78*	73.60
Mean Random	.54	.30	75.17*	75.44
Best Three	.71	.49	167.57*	64.03

Table A-3
Four-Engagement Random Subsets vs The Four Best Predictors

Engagements	Multiple R	Adjusted R^2	$F(1,169)$	SE
A1, B5, B2/B2A, A5/A5A	.60	.36	94.23*	72.33
B1, B3S, B5, A5/A5A	.68	.46	145.08*	66.29
A4, A3, B4, B2/B2A	.63	.40	113.99*	69.83
A2S. B3S. B5, A3	.63	.40	112.64*	70.00
B1, B2/B2A, A4, A1	.66	.44	133.15*	67.58
Mean Random	.64	.41	120.94	69.21
Best Four	.77	.59	249.22*	57.45

Five-Engagement random subsets. The first five rows of Table A-4 present the shortcut-method-based multiple regression results for the five randomly drawn subsets containing five engagements. Means in the sixth line of the table are based on these subsets. The last line in the table provides analogous multiple regression results based on the four best predictors (Engagements B1, A2S, A3, A5/A5A, and B2/B2A). Adjusted R^2s for the random subsets ranged from .45 to .60 and accounted for an average of 53% of criterion variance and produced an average *SE* of 61.47 along with significant *F*-values in excess of 139. By comparison, the two best predictors accounted for 68% of adjusted criterion variance and an *F*-value of 366.86.

Six-Engagement random subsets. The first five rows of Table A-5 present the shortcut-method-based multiple regression results for random subsets containing six engagements. Means in the sixth line of the table are based on these subsets. The last line in the table provides analogous multiple regression results based on the four best predictors (Engagements B1, A2S, A3, A5/A5A, B2/B2A, and A1). Adjusted R^2s for the random subsets ranged from .49 to .67 and accounted for an average of 60% of criterion variance and produced an average *SE* of 57.03 along with significant *F*-values in excess of 167. By comparison, the two best predictors accounted for 76% of adjusted criterion variance and an *F*-value of 534.77.

Table A-4
Five-Engagement Random Subsets vs The Five Best Predictors

Engagements	Multiple R	Adjusted R^2	$F(1,169)$	SE
A3, A4, A5/A5A, B1, B3S	.73	.53	195.73*	61.51
A1, A2S, A5/A5A, B1, B5	.77	.60	253.46*	57.16
A3, A4, A5/A5A, B1, B4	.73	.53	194.74*	61.60
A2S,A5/A5A,B2/B2A,B4,B5	.67	.45	139.40*	66.90
A1, A2S, A3, A4, B2/B2A	.75	.55	211.84*	60.20
Mean Random	.73	.53	199.03	61.47
Best 5	.83	.68	366.86*	50.75

Table A-5
Six-Engagement Random Subsets vs The Six Best Predictors

Excluded Engagements	Multiple R	Adjusted R^2	$F(1,169)$	SE
A1, A3, B3S, B5	.82	.67	351.36*	51.50
A3, A5/A5A, B3S, B5	.81	.66	327.48*	52.72
A3, A4, B1, B3S	.75	.56	213.90*	60.04
A2S, B1, B2/B2A, B5	.71	.49	167.52*	64.04
A2S, A4, B2S, B5	.78	.60	257.71*	56.87
Mean Random	.77	.60	263.59	57.03
Best Six	.87	.76	534.77*	44.28

Seven-engagement random subsets. The first five rows of Table A-6 present the shortcut-method-based multiple regression results for the five randomly drawn subsets containing seven engagements. Means in the sixth line of the table are based on these subsets. The last line in the table provides analogous multiple regression results based on the four best predictors (Engagements B1, A2S, A3, A5/A5A, B2/B2A, A1, and A4). Adjusted R^2s for the random subsets ranged from .66 to .81 and accounted for an average of 72% of criterion variance and produced an average *SE* of 47.70 along with significant *F*-values in excess of 327. By comparison, the two best predictors accounted for 84% of adjusted criterion variance and had an F-value of 892.13.

Eight-Engagement random subsets. The first five rows of Table A-7 present the shortcut-method-based multiple regression results for the five randomly drawn subsets containing eight engagements. Means in the sixth line of the table are based on these subsets. The last line in the table provides analogous multiple regression results based on the four best predictors (Engagements B1, A2S, A3, A5/A5A, B2/B2A, A1, A4, and B4). Adjusted R^2s for the random subsets ranged from .71 to .90 and accounted for an average of 82% of criterion variance and produced an average *SE* of 37.25 along with significant *F*-values in excess of 424. By comparison, the two best predictors accounted for 90% of adjusted criterion variance and had an F-value of 1471.52.

Table A-6
Seven-Engagement Random Subsets vs The Seven Best Predictors

Excluded Engagements	Multiple R	Adjusted R^2	*F*(1,169)	SE
A1, A5/A5A, B2/B2A	.84	.71	414.50*	48.63
A5/A5A, B1, B5	.81	.66	327.69*	52.71
A2S, B2/B2A, B4	.83	.69	382.04*	50.05
A3, A4, B3S	.85	.71	432.37*	47.91
A4, B4, B5	.90	.81	730.06*	39.18
Mean Random	.85	.72	457.33	47.70
Best Seven	.92	.84	892.13*	36.06

Table A-7
Eight-Engagement Random Subsets vs The Eight Best Predictors

Excluded Engagements	Multiple R	Adjusted R^2	*F*(1,169)	SE
B4, B5	.94	.88	1249.43*	31.19
A1, A4	.91	.82	818.51*	37.38
B3S, B5	.95	.90	1471.52*	29.00
A3, B1	.85	.71	424.78*	48.21
A1, B2/B2A	.89	.80	673.40*	40.48
Mean Random	.91	.82	927.53	37.25
Best Eight	.95	.90	1471.52*	29.00

Nine-Engagements random subsets. The first five rows of Table A-8 present the shortcut-method-based multiple regression results for the five randomly drawn subsets containing nine engagements. Means in the sixth line of the table are based on these subsets. The last line in the table provides analogous multiple regression results based on the four best predictors (Engagements B1, A2S, A3, A5/A5A, B2/B2A, A1, A4, B4, and B3S). Adjusted R^2s for the random subsets ranged from .82 to .94 and accounted for an average of 89% of criterion variance and produced an average *SE* of 29.12 along with significant *F*-values in excess of 755. By comparison, the two best predictors accounted for 95% of adjusted criterion variance and had an F-value of 3145.88.

Table A-8
Nine-Engagement Random Subset vs The Nine Best Predictors

Excluded Engagements	Multiple R	Adjusted R^2	$F(1,169)$	SE
A2S	.94	.89	1347.80*	30.16
B3S	.97	.94	2708.85*	21.90
B2/B2A	.95	.90	1612.65*	27.83
B1	.90	.82	755.10*	38.65
A4	.95	.91	1716.22*	27.06
Mean Random	.94	.89	1628.12	29.12
Best Nine	.97	.95	3145.88*	20.40